0 Preliminaries

The main ingredients of snacks are sugar and fat. The main ingredients of math are logic and set theory.

0.1 First-order logic

Definition 0.1. A set \( S \) is a collection of distinct objects \( x \)'s, often denoted with the following notation
\[
S = \{ x \mid \text{the conditions that } x \text{ satisfies.} \}.
\]

Notation 1. \( \mathbb{R}, \mathbb{Z}, \mathbb{N}, \mathbb{Q}, \mathbb{C} \) denote the sets of real numbers, integers, natural numbers, rational numbers and complex numbers, respectively. \( \mathbb{R}^+, \mathbb{Z}^+, \mathbb{N}^+, \mathbb{Q}^+ \) the sets of positive such numbers.

Definition 0.2. \( S \) is a subset of \( \mathcal{U} \), written as \( S \subseteq \mathcal{U} \), if and only if (iff) \( x \in S \Rightarrow x \in \mathcal{U} \). \( S \) is a proper subset of \( \mathcal{U} \), written as \( S \subset \mathcal{U} \) and \( \exists x \in \mathcal{U} \) s.t. \( x \notin S \).

Definition 0.3 (Statements of first-order logic). A universal statement is a logic statement of the form
\[
U = (\forall x \in S, A(x)).
\]
An existential statement has the form
\[
E = (\exists x \in S, \text{s.t. } A(x)),
\]
where \( \forall \) ("for each") and \( \exists \) ("there exists") are the quantifiers, \( S \) is a set, "s.t." means "such that," and \( A(x) \) is the formula.
A statement of implication/conditional has the form
\[
A \Rightarrow B.
\]

Example 0.1. universal and existential statements:
\[
\forall x \in [2, +\infty), x > 1; \\
\forall x \in \mathbb{R}^+, x > 1; \\
\exists p, q \in \mathbb{Z}, \text{s.t. } p/q = \sqrt{2}; \\
\exists p, q \in \mathbb{Z}, \text{s.t. } \sqrt{p} = \sqrt{q} + 1;
\]

Remark 0.2. A logic statement is either true or false. There is no such thing that a logic statement is sometimes true and sometimes false. To prove a universal statement, conceptually we have to verify the statement for all elements in the set. To deny a universal statement, we only need to show a counterexample. To prove an existential statement, we only need to show an instance. To deny an existential statement, conceptually we have to show that the statement holds for none of the elements.

Remark 0.3. In Definition 0.3, the formula \( A(x) \) itself might also be a logic statement. Hence universal and existential statements might be nested. This observation leads to the next definition.

Definition 0.4. A universal-existential statement is a logic statement of the form
\[
U_E = (\forall x \in S, \exists y \in T \text{ s.t. } A(x, y)).
\]
An existential-universal statement has the form
\[
E_U = (\exists y \in T, \text{s.t. } \forall x \in S, A(x, y)).
\]

Example 0.4. True or false:
\[
\forall x \in [2, +\infty), \exists y \in \mathbb{Z}^+ \text{ s.t. } x^y < 10^3; \\
\exists y \in \mathbb{R} \text{ s.t. } \forall x \in [2, +\infty), x > y; \\
\exists y \in \mathbb{R} \text{ s.t. } \forall x \in [2, +\infty), x < y;
\]

Example 0.5 (Translating a English statement into a logic statement). Goldbach's conjecture states every even natural number greater than 2 is the sum of two primes. Let \( \mathbb{P} \subset \mathbb{N}^+ \) denote the set of prime numbers. Then Goldbach's conjecture is \( \forall a \in 2\mathbb{N}^+ + 2, \exists p, q \in \mathbb{P}, \text{s.t. } a = p + q \).

Theorem 0.5. The existential-universal statement implies the corresponding universal-existential statement, but not vice versa.

Example 0.6 (Translating a logic statement to an English statement). Let \( S \) be the set of all human beings. \( U_E = (\forall p \in S, \exists q \in S \text{ s.t. } q \text{ is } p \text{'s mom.}) \\
E_U = (\exists q \in S \text{ s.t. } \forall p \in S, \text{ } p \text{'s mom.}) \\
U_E \) is probably true, but \( E_U \) is certainly false. If \( E_U \) were true, then \( U_E \) would be true. why?

Axiom 0.6 (First-order negation of logical statements). The negations of the statements in Definition 0.3 are
\[
\neg U = (\exists x \in S, \text{ s.t. } \neg A(x)).
\]
\[
\neg E = (\forall x \in S, \neg A(x)).
\]

Example 0.7. The negation of a more complicated logic statement abides by the following rules:
- switch the type of each quantifier until you reach the last formula without quantifiers;
- negate the last formula.

One might need to group quantifiers of like type.

Example 0.8 (The negation of Goldbach's conjecture). \( \exists a \in 2\mathbb{N}^+ + 2 \text{ s.t. } \forall p, q \in \mathbb{P}, a \neq p + q \).

This conjecture has been shown to hold up through \( 4 \times 10^{18} \), but no proofs and disproofs have been found.

Example 0.9. Negation of the statement in Definition 0.49.

Axiom 0.7 (Contraposition). A conditional statement is logically equivalent to its contrapositive.
\[
(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)
\]

Example 0.10. "If Jack is a man, then Jack is a human being," is equivalent to "If Jack is not a human being, then Jack is not a man."

Exercise 0.11. Draw an Euler diagram of subsets to illustrate Example 0.10.
0.2 Ordered Sets

**Definition 0.8.** The Cartesian product $\mathcal{X} \times \mathcal{Y}$ between two sets $\mathcal{X}$ and $\mathcal{Y}$ is the set of all possible ordered pairs with first element from $\mathcal{X}$ and second element from $\mathcal{Y}$:

$$\mathcal{X} \times \mathcal{Y} = \{(x, y) \mid x \in \mathcal{X},\ y \in \mathcal{Y}\}. \quad (0.10)$$

**Axiom 0.9** (Fundamental principle of counting). A task consists of a sequence of $k$ steps. Let $n_i$ denote the number of different choices for the $i$-th step, the total number of distinct ways to complete the task is then

$$\prod_{i=1}^{k} n_i = n_1 n_2 \cdots n_k. \quad (0.11)$$

**Example 0.12.** Let $A, E, D$ be the set of appetizers, main entrees, desserts in a restaurant. $A \times E \times D$ is the set of possible dinner combos. If $\#A = 10, \#E = 5, \#D = 6, \#(A \times E \times D) = 300$.

**Definition 0.10** (Maximum and minimum). Consider $S \subseteq \mathbb{R}, S \neq \emptyset$. If $\exists m \in S$ s.t. $\forall x \in S, x \leq m$, then $m$ is the maximum of $S$ and denoted by $\max S$. If $\exists m \in S$ s.t. $\forall x \in S, x \geq m$, then $m$ is the minimum of $S$ and denoted by $\min S$.

**Definition 0.11** (Upper and lower bounds). Consider $S \subseteq \mathbb{R}, S \neq \emptyset$. $a$ is an upper bound of $S \subseteq \mathbb{R}$ if $\forall x \in S, x \leq a$; then the set $S$ is said to be bounded above. $a$ is a lower bound of $S$ if $\forall x \in S, x \geq a$; then the set $S$ is said to be bounded below. $S$ is bounded if it is bounded above and bounded below.

One difference between a maximum and an upper bound is that the former belongs to the set while the latter might not.

**Definition 0.12** (Supremum and infimum). Consider $S \subseteq \mathbb{R}, S \neq \emptyset$. If $S$ is bounded above and $S$ has a least upper bound then we call it the supremum of $S$ and denote it by $\sup S$. If $S$ is bounded below and $S$ has a greatest lower bound, then we call it the infimum of $S$ and denote it by $\inf S$.

**Example 0.13.** If $S$ has a maximum, then $\max S = \sup S$.

**Example 0.14.** $\sup [a, b] = \sup\{a, b\} = \sup\{a, b\} = \sup(a, b)$.

**Axiom 0.13** (Completeness of $\mathbb{R}$). Every nonempty subset of $\mathbb{R}$ that is bounded above has a least upper bound.

In other words, for any nonempty $S \subseteq \mathbb{R}$ bounded above, $\sup S$ exists and is a real number.

**Corollary 0.14.** Every nonempty subset of $\mathbb{R}$ that is bounded below has a greatest lower bound.

**Definition 0.15.** A binary relation between two sets $\mathcal{X}$ and $\mathcal{Y}$ is an ordered triple $(\mathcal{X}, \mathcal{Y}, g)$ where $g \subseteq \mathcal{X} \times \mathcal{Y}$.

A binary relation on $\mathcal{X}$ is the relation between $\mathcal{X}$ and $\mathcal{X}$.

The statement $(x, y) \in R$ is read “$x$ is $R$-related to $y$,” and denoted by $x R y$ or $R(x, y)$.

**Definition 0.16.** A binary relation “$\leq$” on some set $S$ is a total order or linear order on $S$ iff, $\forall a, b, c \in S$,

- $a \leq b$ and $b \leq a$ imply $a = b$ (antisymmetry);
- $a \leq b$ and $b \leq c$ imply $a \leq c$ (transitivity);
- $a \leq b$ or $b \leq a$ (totality).

A set equipped with a total order is a chain or totally ordered set.

**Example 0.15.** The real numbers with less or equal.

**Example 0.16.** The English letters of the alphabet with dictionary order.

**Example 0.17.** The Cartesian product of a set of totally ordered sets with the lexicographical order.

**Example 0.18.** Sort your book in lexicographical order and save a lot of time. $\log_{26} N \ll N$!

**Definition 0.17.** A binary relation “$\leq$” on some set $S$ is a partial order on $S$ iff, $\forall a, b, c \in S$, antisymmetry, transitivity, and reflexivity ($a \leq a$) hold.

A set equipped with a partial order is a called a poset.

**Example 0.19.** The set of subsets of a set $S$ ordered by inclusion “$\subseteq$.”

**Example 0.20.** The natural numbers equipped with the relation of divisibility.

**Example 0.21.** The set of stuff you will put on your body every morning with the time ordered: undershorts, pants, belt, shirt, tie, jacket, socks, shoes, watch.

**Example 0.22.** Inheritance (“is-a” relation) is a partial order. $A \rightarrow B$ reads “$B$ is a special type of $A$.”

**Example 0.23.** Composition (“has-a” relation) is also a partial order. $A \mapsto B$ reads “$B$ has an instance/object of $A$.”

**Example 0.24.** Implication “$\Rightarrow$” is a partial order on the set of logical statements.

**Example 0.25.** The set of definitions, axioms, propositions, theorems, lemma, etc. is a poset with inheritance, composition, and implication. It is helpful to relate them with these partial orderings.

“If syntax sugar does not count, there is nothing left.”

0.3 Functions: limits and continuity

**Definition 0.18.** A function/map/mapping $f$ from $\mathcal{X}$ to $\mathcal{Y}$, written as $f : \mathcal{X} \rightarrow \mathcal{Y}$ or $\mathcal{X} \rightarrow \mathcal{Y}$, is a subset of the Cartesian product $\mathcal{X} \times \mathcal{Y}$ satisfying that $\forall x \in \mathcal{X}$, there is exactly one $y \in \mathcal{Y}$ s.t. $(x, y) \in \mathcal{X} \times \mathcal{Y}$. $\mathcal{X}$ and $\mathcal{Y}$ are the domain and range of $f$, respectively.

**Remark 0.26.** The important thing in the above definition is the uniqueness of the pair $(x, y)$. Why?
Definition 0.19. A function \( f : \mathcal{X} \to \mathcal{Y} \) is said to be injective or one-to-one if
\[
\forall x_1 \in \mathcal{X}, \forall x_2 \in \mathcal{X}, \ x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2).
\] (0.12)
It is surjective or onto if
\[
\forall y \in \mathcal{Y}, \exists x \in \mathcal{X}, \ s.t. \ y = f(x).
\] (0.13)
It is bijective iff it is both injective and surjective.

Definition 0.20. A set \( S \) is countably infinite iff there exists a bijective function \( f : S \to \mathbb{N}^+ \) that maps \( S \) to \( \mathbb{N}^+ \). A set is countable if it is either finite or countably infinite.

Example 0.27. Is the integers countable? Is the rationals countable? Is the real numbers countable?

Definition 0.21. A scalar function is a function whose range is a subset of \( \mathbb{R} \).

Definition 0.22 (Limit of a scalar function with one variable). Consider a function \( f : I \to \mathbb{R} \) with \( I(c, r) = (c-r, c) \cup (c, c+r) \). The limit of \( f(x) \) exists as \( x \) approaches \( c \) written as \( \lim_{x \to c} f(x) = L \), iff
\[
\forall \epsilon > 0, \exists \delta > 0, \ s.t. \ \forall x \in (c, \delta), \ |f(x) - L| < \epsilon.
\] (0.14)
The notation reads "as \( x \) gets closer to \( c \), \( f(x) \) gets closer to \( L \)." How close is close? As close as you wish. This idea is packaged in the \( \epsilon - \delta \) technique.

Example 0.28. show that \( \lim_{x \to 2} \frac{1}{x} = \frac{1}{2} \).

Proof. If \( \epsilon \geq \frac{1}{2} \), choose \( \delta = 1 \). Then \( x \in (1,3) \) implies \( |\frac{1}{2} - 1| < \frac{1}{2} \) since \( \frac{1}{2} - 1 \) is a monotonically decreasing function with \( \frac{1}{2} \) as its supremum at \( x = 1 \).

If \( \epsilon \in (0, \frac{1}{2}) \), choose \( \delta = \epsilon \). Then \( x \in (2-\epsilon, 2+\epsilon) \subset \left( \frac{3}{2}, \frac{5}{2} \right) \). Hence \( \frac{1}{2} - 1 = \frac{1}{2} \left( \frac{1}{2} - 1 \right) < |2-x| < \epsilon \). The proof is completed by Definition 0.22.

The philosophy here is that if you can make the difference of two functions as small as you wish, then they have the same limit.

Definition 0.23. \( f : \mathbb{R} \to \mathbb{R} \) is continuous at \( c \) iff
\[
\lim_{x \to c} f(x) = f(c).
\] (0.15)
\( f \) is continuous on \( (a,b) \), written as \( f \in C(a,b) \) if (0.15) holds \( \forall x \in (a,b) \).

Definition 0.24. Let \( I = (a,b) \). A function \( f : I \to \mathbb{R} \) is uniformly continuous on \( I \) iff
\[
\forall \epsilon > 0, \exists \delta > 0, \ s.t. \ \forall x, y \in I, \ |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.
\] (0.16)

Example 0.29. On \( (a, \infty) \), \( f(x) = \frac{1}{x} \) is uniformly continuous if \( a > 0 \) and is not so if \( a = 0 \).

Proof. If \( a > 0 \), then \( |f(x) - f(y)| = |\frac{x-y}{xy}| < |\frac{x-y}{a^2}| \).
Hence \( \forall \epsilon > 0, \exists \delta = a^2 \epsilon, \ s.t. \ |x-y| < \delta \Rightarrow |f(x) - f(y)| < \frac{|x-y|}{a^2} < \frac{x^2}{a^2} = \epsilon \).

If \( a = 0 \), negating the condition of uniform continuity,
\( |f(x) - f(y)| = \frac{|x-y|}{xy} \) or \( \epsilon \). We prove a stronger version: \( \forall \epsilon > 0, \forall \delta > 0 \exists x, y > 0 \) s.t. \( |x - y| < \delta \Rightarrow |f(x) - f(y)| \geq \epsilon \).
We prove a stronger version: \( \forall \epsilon > 0, \forall \delta > 0 \exists x, y > 0 \) s.t. \( |x - y| < \delta \Rightarrow |f(x) - f(y)| \geq \epsilon \).

Exercise 0.30. On \( (a, \infty) \), \( f(x) = \frac{1}{x^2} \) is uniformly continuous if \( a > 0 \) and is not so if \( a = 0 \).

Theorem 0.25. Uniform continuity implies continuity but the converse is not true.

Proof.

Theorem 0.26. \( f : \mathbb{R} \to \mathbb{R} \) is uniformly continuous on \( (a,b) \) iff it can be extended to a continuous function \( \hat{f} \) on \([a,b]\).

Definition 0.27. The derivative of a function \( f : \mathbb{R} \to \mathbb{R} \) at \( a \) is the limit
\[
f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.
\] (0.17)
If the limit exists, \( f \) is differentiable at \( a \).

Example 0.31. The derivative of power function \( f(x) = x^\alpha \) is \( f' = \alpha x^{\alpha - 1} \).

Proof. use Newton's generalized binomial theorem.
\[
(a + h)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} a^{\alpha-n}h^n.
\]
Think about the three cases \( \alpha = \frac{1}{2}, \alpha = 1, \alpha = 2 \). If \( x \) is time and \( f(x) \) measures how knowledgeable you are. You probably want to have \( f(x) = x^2 \) rather than \( f(x) = x^\frac{1}{2} \). The reason \( f(x) = x^2 \) is much better is that its rate of increase also increase.

Definition 0.28. A function \( f(x) \) is \( k \) times continuously differentiable on \((a,b)\) iff \( f^{(k)}(x) \) exists on \((a,b)\) and is itself continuous. The set or space of all such functions on \((a,b)\) is denoted by \( C^k(a,b) \).
In comparison, \( C^k[a,b] \) is the space of functions \( f \) that \( f^{(k)}(x) \) is bounded and uniformly continuous on \((a,b)\).

Theorem 0.29. A scalar function \( f \) is bounded on \([a,b]\) if \( f \in C^k[a,b] \).

Theorem 0.30 (Intermediate value). A scalar function \( f \in C[a,b] \) satisfies
\[
\forall y \in [m,M], \exists \xi \in [a,b], \ s.t. \ y = f(\xi)
\] (0.18)
where \( m = \inf_{x \in [a,b]} f(x) \) and \( M = \sup_{x \in [a,b]} f(x) \).

This is a more explicit version of that in the book, which states that a continuous function assumes all values between \( f(a) \) and \( f(b) \) on a closed interval \([a, b]\).
Theorem 0.31. If \( f : (a, b) \to \mathbb{R} \) assumes its maximum or minimum at \( x_0 \in (a, b) \) and \( f \) is differentiable at \( x_0 \), then \( f'(x_0) = 0 \).

Proof. Suppose \( f'(x_0) > 0 \), then \( f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} > 0 \). By the definition of a limit, \( \exists \delta > 0 \) s.t. \( a < x_0 - \delta < x_0 + \delta < b \) and \( |x - x_0| < \delta \) imply \( \frac{f(x) - f(x_0)}{x - x_0} > 0 \), which is a contradiction to \( f(x_0) \) being a maximum when we choose \( x \in (x_0, x_0 + \delta) \).

Theorem 0.32 (Rolle’s). If a function \( f : \mathbb{R} \to \mathbb{R} \) satisfies

(i) \( f \in C[a, b] \) and \( f' \) exists on \( (a, b) \),

(ii) \( f(a) = f(b) \),

then \( \exists x \in (a, b) \) s.t. \( f'(x) = 0 \).

Proof. By Theorem 0.30, all values between sup \( f \) and inf \( f \) will be assumed. If \( f(a) = f(b) = \sup f = \inf f \), then \( f \) is a constant on \([a, b]\) and thus the conclusion holds. Otherwise, Theorem 0.31 completes the proof.

Example 0.32. Whether the sequence starts from 0 or 1 is a matter of convention and convenience according to the context.

Definition 0.35 (Limit of a sequence). A sequence \( \{a_n\} \) has the limit \( L \), written as \( \lim_{n \to \infty} a_n = L \), or \( a_n \to L \) as \( n \to \infty \), iff

\[
\forall \varepsilon > 0, \exists N, \text{ s.t. } \forall n > N, \ |a_n - L| < \varepsilon. \tag{0.19}
\]

If such a limit \( L \) exists, we say that \( \{a_n\} \) converges to \( L \).

Example 0.33 (A sequence can be used to approximate the limit with increasing accuracy: the story of \( \pi \)).

\[
\pi \approx 3.141592653589793. \tag{0.20}
\]

Real world applications often require that we approximate a number or function with varying accuracy depending on the scenarios. This is one of the main motivations of infinite sequence and series.

\begin{itemize}
  \item In NASA, calculations involving \( \pi \) use 15 digits for Guidance Navigation and Control;
  \item if you want to compute the circumference of the entire universe to the accuracy of less than the diameter of a hydrogen atom, you need only 39 decimal places of \( \pi \).
\end{itemize}

We human beings have more than a trillion digits. A very good estimate is given by Zu, ChongZhi 1500 years ago:

\[
\pi \approx \frac{355}{113} = 3.14159292, \tag{0.21}
\]

which is good enough for estimating the length of the perimeter of your backyard foundation.

Theorem 0.36 (Bolzano-Weierstrass). Every bounded sequence has a convergent subsequence.

Definition 0.37. A series associated with an infinite sequence \( \{a_n\} \) is defined as \( \sum_{n=0}^{\infty} a_n \), the sum of all terms of the sequence.

Definition 0.38. The sequence of partial sums \( S_n \) associated to a series \( \sum_{n=0}^{\infty} a_n \) is defined for each \( n \) as the sum of the sequence \( \{a_i\} \) from \( a_0 \) to \( a_n \):

\[
S_n = \sum_{i=0}^{n} a_i. \tag{0.22}
\]

Proposition 0.39. A series converges to \( L \) iff the associated sequence of partial sums converges to \( L \).

Definition 0.40. A power series centered at \( c \) is a series of the form

\[
p(x) = \sum_{n=0}^{\infty} a_n(x-c)^n, \tag{0.23}
\]

where \( a_n \)'s are the coefficients. The interval of convergence is the set of values of \( x \) for which the series converges:

\[
I_c(p) = \{x \mid p(x) \text{ converges}\}. \tag{0.24}
\]

Definition 0.41. If the derivatives \( f^{(i)}(x) \) with \( i = 1, 2, \ldots, n \) exist for a function \( f : \mathbb{R} \to \mathbb{R} \) at \( x = c \), then

\[
T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!} (x-c)^k, \tag{0.25}
\]

is called the \( n \)th Taylor polynomial for \( f(x) \) at \( c \). In particular, the linear approximation for \( f(x) \) at \( c \) is

\[
T_1(x) = f(c) + f'(c)(x-c). \tag{0.26}
\]

Example 0.34. If \( f \in C^\infty \), then \( \forall n \in \mathbb{N} \),

\[
T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!} (x-c)^k, \tag{0.27}
\]

\begin{itemize}
  \item if \( f \in C^\infty \), then \( \forall n \in \mathbb{N} \),
  \item \( T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!} (x-c)^k \). \tag{0.28}
\end{itemize}

Proof. By induction. In the inductive step, regroup the summation into a constant term and another shifted summation.

Definition 0.42. The Taylor series (or Taylor expansion) for \( f(x) \) at \( c \) is

\[
\sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!} (x-c)^k. \tag{0.29}
\]
There exists some $c$ and $\delta > 0$ between $R$ and $f \in C^\infty \rightarrow C$. If (Taylor’s theorem with Lagrangian form)

$$\lim_{n \rightarrow \infty} E_n(x) = 0 \iff \lim_{n \rightarrow \infty} T_n(x) = f(x).$$

Let $T_n$ be the $n$th Taylor polynomial for $f(x)$ at $c$.

**Theorem 0.45.**

**Lemma 0.45.**

$\forall m = 0, 1, 2, \ldots, n, E^{(m)}(c) = 0$.

**Theorem 0.46** (Taylor’s theorem with Lagrangian form). Consider a function $f : \mathbb{R} \rightarrow \mathbb{R}$. If $f \in C^n [c - d, c + d]$ and $f^{(n+1)}(x)$ exists on $(c - d, c + d)$, then $\forall x \in [c - d, c + d]$, there exists some $\xi$ between $c$ and $x$ such that

$$E_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - c)^{n+1}.$$ 

**Proof.** Fix $x \neq c$, let $M$ be the unique solution of

$$E_n(x) = f(x) - T_n(x) = \frac{M(x - c)^{n+1}}{(n+1)!}.$$ 

Consider function

$$g(t) = E_n(t) - \frac{M(t - c)^{n+1}}{(n+1)!}.$$ 

Clearly $g(x) = 0$. By Lemma 0.45, $g^{(k)}(c) = 0$ for $k = 0, 1, \ldots, n$. Then Rolle’s theorem implies that

$$\exists x_1 \in (c, x) \text{ s.t. } g'(x_1) = 0.$$ 

If $x < c$, change $(c, x)$ above to $(x, c)$. Apply Rolle’s theorem to $g'(t)$ on $(c, x_1)$ and we have

$$\exists x_2 \in (c, x_1) \text{ s.t. } g''(x_2) = 0.$$ 

Repeatedly using Rolle’s theorem,

$$\exists x_{n+1} \in (c, x_n) \text{ s.t. } g^{(n+1)}(x_{n+1}) = 0.$$ 

Since $T_n^{(n+1)}(t) = 0$, $g^{(n+1)}(t) = M - f^{(n+1)}(t) = 0$, hence $M = f^{(n+1)}(x_{n+1})$. The proof is completed by identifying $\xi$ with $x_{n+1}$. 

**Example 0.35.** How many terms are needed to compute $e^2$ correctly to four decimal places?

By D0.42, the Taylor series of $e^x$ at $c = 0$ is

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$ 

By T0.46,

$$\exists \xi \in [0, 2] \text{ s.t. } E_n(2) = e^2\xi^{n+1}/(n+1)! < e^2\xi^{n+1}/(n+1)!.$$ 

Plug in the error criteria and we have $n = 12$, hence 13 terms.

### 0.5 Riemann integral

**Definition 0.47.** A partition of an interval $I = [a, b]$ is a finite ordered subset $T_n \subseteq I$ of the form

$$T_n(a, b) = \{a = x_0 < x_1 < \cdots < x_n = b\}.$$ 

The interval $I_i = [x_{i-1}, x_i]$ is the $i$th subinterval of the partition. The norm of the partition is the length of the longest subinterval,

$$h_n = h(T_n) = \max(x_i - x_{i-1}), \quad i = 1, 2, \ldots, n.$$ 

**Definition 0.48.** The Riemann sum of $f : \mathbb{R} \rightarrow \mathbb{R}$ over a partition $I$ is

$$S_n(f) = \sum_{i=1}^{n} f(x_i^*)(x_i - x_{i-1}).$$ 

where $x_i^* \in I_i$ is a sample point of the $i$th subinterval.

**Definition 0.49.** $f : \mathbb{R} \rightarrow \mathbb{R}$ is integrable (or more precisely Riemann integrable) on $[a, b]$ iff

$$\forall L \in \mathbb{R}, \exists \varepsilon > 0, \exists \delta > 0 \text{ s.t. } \forall I_n(a, b) \text{ with } h(T_n) < \delta, \quad |S_n(f) - L| < \varepsilon.$$ 

**Example 0.36.** The following function $f : [a, b] \rightarrow \mathbb{R}$ is not Riemann integrable.

$$f(x) = \begin{cases} 1 & x \text{ is rational;} \\ 0 & x \text{ is irrational.} \end{cases}$$ 

**Proof.** Negate the equation in Definition 0.49.

$$\forall L \in \mathbb{R}, \exists \varepsilon > 0, \exists \delta > 0 \text{ s.t. } \exists I_n(a, b) \text{ with } h(T_n) < \delta, \quad |S_n(f) - L| \geq \varepsilon.$$ 

For any given $L$, one can instantiate $\varepsilon = \frac{L - a}{b - a}$ such that, no matter how small $\delta$ is, one can choose $x_i^* \in S_n(f)$ such that the distance between $S_n(f)$ and $L$ is greater than $\varepsilon$.

**Definition 0.50.** If $f : \mathbb{R} \rightarrow \mathbb{R}$ is integrable on $[a, b]$, then its limit is called the definite integral of $f$ on $[a, b]$:

$$\int_{a}^{b} f(x)dx = \lim_{h_n \rightarrow 0} S_n(f).$$ 

**Theorem 0.51.** A scalar function $f$ is integrable on $[a, b]$ if $f \in C[a, b]$.

**Definition 0.52.** A monotonic function is a function between ordered sets that either preserves or reverses the given order. In particular, $f : \mathbb{R} \rightarrow \mathbb{R}$ is monotonically increasing if $\forall x, y, x \leq y \Rightarrow f(x) \leq f(y)$; $f : \mathbb{R} \rightarrow \mathbb{R}$ is monotonically decreasing if $\forall x, y, x \leq y \Rightarrow f(x) \geq f(y)$.

Here ordered sets can be posets, but we will limit our attention to chains.

**Theorem 0.53.** A scalar function is integrable on $[a, b]$ if it is monotonic on $[a, b]$. 

Qinghai Zhang Numerical Approximation 2018
**Example 0.37.** True or false: a bijective function is either order-preserving or order-reversing?

False; missing continuity. In other words, a continuous bijective function is either order-preserving or order-reversing.

**Theorem 0.54 (Integral mean value).** Let \( w : [a, b] \to \mathbb{R}^+ \) be integrable on \([a, b]\). For \( f \in \mathcal{C}[a, b] \), \( \exists \xi \in [a, b] \) s.t.

\[
\int_a^b w(x)f(x)\,dx = f(\xi)\int_a^b w(x)\,dx.
\]

(0.34)

**Proof.** Denote \( m = \inf_{x\in[a,b]} f(x) \), \( M = \sup_{x\in[a,b]} f(x) \), and \( I = \int_a^b w(x)\,dx \). Then \( mw(x) \leq f(x)w(x) \leq Mw(x) \) and

\[
mI \leq \int_a^b w(x)f(x)\,dx \leq MI.
\]

If \( w > 0 \) implies \( I \neq 0 \), hence

\[
m \leq \frac{1}{I} \int_a^b w(x)f(x)\,dx \leq M.
\]

Applying Theorem 0.30 completes the proof. \( \square \)

### 0.6 Vector spaces

**Definition 0.55.** A field is a commutative division ring.

More commonly, a field \( \mathbb{F} \) is a set together with two binary operations, usually called “addition” and “multiplication” and denoted by “+” and “∗”, such that \( \forall a, b, c \in \mathbb{F} \), the following axioms hold,

- commutativity: \( a + b = b + a, \ ab = ba \);
- associativity: \( a + (b + c) = (a + b) + c, \ a(bc) = (ab)c \);
- identity: \( a + 0 = a, \ a1 = a \);
- invertibility: \( a + (−a) = 0, \ aa^{-1} = 1 \) (\( a \neq 0 \));
- distributivity: \( a(b + c) = ab + ac \).

**Definition 0.56.** A vector space or linear space over a field \( \mathbb{F} \) is a set \( \mathcal{V} \) together with two binary operations “+” and “∗” respectively called vector addition and scalar multiplication that satisfy the following axioms:

(VSA-1) commutativity

\( \forall u, v \in \mathcal{V}, \ u + v = v + u \);

(VSA-2) associativity

\( \forall u, v, w \in \mathcal{V}, \ (u + v) + w = u + (v + w) \);

(VSA-3) compatibility

\( \forall u \in \mathcal{V}, \forall a, b \in \mathbb{F}, \ (ab)u = a(bu) \);

(VSA-4) additive identity

\( \forall u \in \mathcal{V}, \ \exists 0 \in \mathbb{F}, \ \text{s.t.} \ u + 0 = u \);

(VSA-5) additive inverse

\( \forall u \in \mathcal{V}, \ \exists v \in \mathcal{V}, \ \text{s.t.} \ u + v = 0 \);

(VSA-6) multiplicative identity

\( \forall u \in \mathcal{V}, \ \exists 1 \in \mathbb{F}, \ \text{s.t.} \ 1u = u \);

(VSA-7) distributive laws

\[
\forall u, v \in \mathcal{V}, \ \forall a, b \in \mathbb{F}, \ \left\{ \begin{array}{l}
(a + b)u = au + bu, \\
(a(u + v)) = au + av.
\end{array} \right.
\]

The elements of \( \mathcal{V} \) are called vectors and the elements of \( \mathbb{F} \) are called scalars.

**Definition 0.57.** A vector space with \( \mathbb{F} = \mathbb{R} \) or \( \mathbb{F} = \mathbb{C} \) is called a real vector space or a complex vector space, respectively.

**Example 0.38.** We will limit ourselves to these two vector spaces for the whole semester.

**Example 0.39.** The simplest vector space is \( \{ \mathbf{0} \} \). Another simple example of a vector space over a field \( \mathbb{F} \) is itself, equipped with its standard addition and multiplication.

**Definition 0.58.** A list of length \( n \) or \( n \)-tuple is an ordered collection of \( n \) elements (which might be numbers, other lists, or more abstract entities) separated by commas and surrounded by parentheses: \( x = (x_1, x_2, \ldots, x_n) \).

**Example 0.40.** According to the definition, two lists are equal iff they have the same length and the same elements in the same order. Hence the two lists \( (3, 5) \) and \( (5, 3) \) are not equal while the two sets \( \{3, 5\} \) and \( \{5, 3\} \) are equal.

The distinguishing features of a list from a set: finite length, ordering, repetition of elements.

**Example 0.41.** A vector space composed of all the \( n \)-tuples of a field \( \mathbb{F} \) is known as a coordinate space, denoted by \( \mathbb{F}^n \) (\( n \in \mathbb{N}^+ \)).

**Example 0.42.** The properties of forces or velocities in the real world can be captured by a coordinate space \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \).

**Example 0.43.** The set of continuous real-valued functions on the interval \([−1, 1]\) forms a real vector space.

**Example 0.44.** For a set \( S \), define

\[ \mathbb{F}^S = \{ f \mid f : S \to \mathbb{F} \} . \]

Then \( \mathbb{F}^S \) is a vector space. The above notation \( \mathbb{F}^n \) can be thought of as a special case of \( \mathbb{F}^S \) because \( n \) can be regarded as \( \{1, 2, \ldots, n\} \) and an element in \( \mathbb{F} \) as a constant function.

**Definition 0.59.** A linear combination of a list of vectors \( \{v_i\} \) is a vector of the form \( \sum a_i v_i \) where \( a_i \in \mathbb{F} \).

**Example 0.45.** \((17, −4, 2)\) is a linear combination of \((2, 1, −3), (1, −2, 4)\) because

\[
(17, −4, 2) = 6(2, 1, −3) + 5(1, −2, 4).
\]

**Example 0.46.** \((17, −4, 5)\) is not a linear combination of \((2, 1, −3), (1, −2, 4)\) because there do not exist numbers \( a_1, a_2 \) such that

\[
(17, −4, 5) = a_1(2, 1, −3) + a_2(1, −2, 4).
\]

Solving from the first two equations yields \( a_1 = 6, a_2 = 5 \).

**Definition 0.60.** The span of a list of vectors \( \{v_i\} \) is the set of all linear combinations of \( \{v_i\} \),

\[
\text{span}(v_1, v_2, \ldots, v_m) = \left\{ \sum_{i=1}^m a_i v_i \mid a_i \in \mathbb{F} \right\} .
\]

(0.35)

In particular, the span of the empty set is \( \{\mathbf{0}\} \). We say that \( (v_1, v_2, \ldots, v_m) \) spans \( \mathcal{V} \) if \( \mathcal{V} = \text{span}(v_1, v_2, \ldots, v_m) \).

Qinghai Zhang Numerical Approximation 2018
Example 0.47.

\[(17, -4, 2) \in \text{span}((2, 1, -3), (1, -2, 4))\]
\[(17, -4, 5) \notin \text{span}((2, 1, -3), (1, -2, 4))\]

Definition 0.61. A vector space \( \mathcal{V} \) is called finite dimensional if some list of vectors span \( \mathcal{V} \); otherwise it is infinite dimensional.

Example 0.48. Let \( \mathcal{P}_m(\mathbb{F}) \) denote the set of all polynomials with coefficients in \( \mathbb{F} \) and degree at most \( m \),

\[\mathcal{P}_m(\mathbb{F}) = \left\{ p : \mathbb{F} \to \mathbb{F} \mid p(z) = \sum_{i=0}^{m} a_iz^i, a_i \in \mathbb{F} \right\}.\]

Then \( \mathcal{P}_m(\mathbb{F}) \) is a finite-dimensional vector space for each non-negative integer \( m \). The set of all polynomials with coefficients in \( \mathbb{F} \), denoted by \( \mathcal{P}(\mathbb{F}) := \mathcal{P}_{\infty}(\mathbb{F}) \), is infinite-dimensional. Both are subspaces of \( \mathcal{F}^\mathbb{F} \) for \( \mathbb{F} = \mathbb{R} \) or \( \mathbb{C} \).

Definition 0.62. A list of vectors \( (v_1, v_2, \ldots, v_m) \) in \( \mathcal{V} \) is called linearly independent iff

\[a_1v_1 + \cdots + a_mv_m = 0 \Rightarrow a_1 = \cdots = a_m = 0. \quad (0.36)\]

Otherwise the list of vectors is called linearly dependent.

Example 0.49. The empty list is declared to be linearly independent. A list of one vector \( (v) \) is linearly independent iff \( v \neq 0 \). A list of two vectors is linearly independent iff neither vector is a scalar multiple of the other.

Example 0.50. The list \( 1, z, \ldots, z^n \) is linearly independent in \( \mathcal{P}_m(\mathbb{F}) \) for each \( m \in \mathbb{N} \).

Example 0.51. \((2, 3, 1), (1, -1, 2), \) and \((7, 3, 8)\) is linearly dependent in \( \mathbb{R}^3 \) because

\[2(2, 3, 1) + 3(1, -1, 2) + (-1)(7, 3, 8) = (0, 0, 0).\]

Example 0.52. Every list of vectors containing the \( 0 \) vector is linearly dependent.

Lemma 0.63 (Linear dependence lemma). Suppose \( \mathcal{V} = (v_1, v_2, \ldots, v_m) \) is a linearly dependent list in \( \mathcal{V} \). Then there exists \( j \in \{1, 2, \ldots, m\} \) such that

- \( v_j \in \text{span}(v_1, v_2, \ldots, v_{j-1}) \);
- if the \( j \)th term is removed from \( \mathcal{V} \), the span of the remaining list equals \( \text{span}(v_1, v_2, \ldots, v_m) \).

Proposition 0.64. In a finite-dimensional vector space, the length of every linearly independent list of vectors is less than or equal to the length of every spanning list of vectors.

Definition 0.65. A basis of a vector space \( \mathcal{V} \) is a list of vectors in \( \mathcal{V} \) that is linearly independent and spans \( \mathcal{V} \).

Example 0.53. In particular, the list of vectors

\[(1, 0, \ldots, 0)^T, \ (0, 1, 0, \ldots, 0)^T, \ldots, \ (0, \ldots, 0, 1)^T \quad (0.37)\]

is called the standard basis of \( \mathbb{F}^n \).

Example 0.54. A basis of \( \mathcal{P}_m(\mathbb{F}) \) in Example 0.48 is \((z^0, z^1, \ldots, z^m)\).

Proposition 0.66. A list of vectors \((v_1, \ldots, v_n)\) is a basis of \( \mathcal{V} \) iff every vector \( u \in \mathcal{V} \) can be written uniquely as

\[u = \sum_{i=1}^{n} a_i v_i, \quad (0.38)\]

where \( a_i \in \mathbb{F} \).

Proposition 0.67. Every spanning list in a vector space \( \mathcal{V} \) can be reduced to a basis of \( \mathcal{V} \). Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of that vector space.

Definition 0.68. The dimension of a finite-dimensional vector space \( \mathcal{V} \), denoted \( \dim \mathcal{V} \), is the length of any basis of the vector space.

Example 0.55. The above definition is well defined because any two bases of a finite-dimensional vector space have the same length.

Proposition 0.69. If \( \mathcal{V} \) is finite-dimensional, then every spanning list of vectors in \( \mathcal{V} \) with length \( \dim \mathcal{V} \) is a basis of \( \mathcal{V} \).

Proposition 0.70. If \( \mathcal{V} \) is finite-dimensional, then every linearly independent list of vectors in \( \mathcal{V} \) with length \( \dim \mathcal{V} \) is a basis of \( \mathcal{V} \).

0.7 Inner product space

Definition 0.71. Let \( \mathbb{F} \) be the underlying field of a vector space \( \mathcal{V} \). The inner product \( \langle u, v \rangle \) on \( \mathcal{V} \) is a function \( \mathcal{V} \times \mathcal{V} \to \mathbb{F} \) that satisfies

(IP-1) real positivity: \( \forall v \in \mathcal{V}, \langle v, v \rangle \geq 0; \)

(IP-2) definiteness: \( \langle v, v \rangle = 0 \iff v = 0; \)

(IP-3) additivity in the first slot:

\[\forall u, v, w \in \mathcal{V}, \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle ;\]

(IP-4) homogeneity in the first slot:

\[\forall \alpha \in \mathbb{F}, \forall v, w \in \mathcal{V}, \langle \alpha v, w \rangle = \alpha \langle v, w \rangle ;\]

(IP-5) conjugate symmetry: \(\forall v, w \in \mathcal{V}, \langle v, w \rangle = \overline{\langle w, v \rangle} .\)

Remark 0.56. The motivation of defining inner products is to capture notions of length, angle, and so on. With length, we can define the limit of a sequence of vectors so that the idea of “approximation” has a solid footing.

Remark 0.57. Complex conjugate of a complex number \( x = a + ib \) is \( \bar{x} = a - ib \). For any two complex numbers \( x, y, \bar{x} + \bar{y} = \bar{x} + \bar{y}, \overline{xy} = \bar{x}\bar{y} \), and \( x\bar{x} = |x|^2 \). Then we have additivity in the second slot and conjugate homogeneity in the second slot.

Definition 0.72. The Euclidean inner product on \( \mathbb{F}^n \) is

\[\langle v, w \rangle = \sum_{i=1}^{n} v_i \bar{w}_i, \quad (0.39)\]
Example 0.58. An inner product can be defined on the vector space of continuous real-valued functions on the interval $[-1, 1]$ by
\[ \langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx. \]

Definition 0.73. An inner product space is a vector space $V$ equipped with an inner product on $V$.

Definition 0.74. Two vectors $u, v$ are called orthogonal if $\langle u, v \rangle = 0$.

Two vectors $u, v$ are orthogonal iff their inner product is the additive identity of the underlying field.

Definition 0.75. Let $F$ be the underlying field of an inner-product space $V$. The norm induced by the inner product on $V$ is a function $V \to F$:
\[ ||v|| = \sqrt{\langle v, v \rangle}. \]
(0.40)

Definition 0.76. The Euclidean $\ell_p$ norm of a vector $v \in F^n$ is
\[ ||v||_p = \left( \sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}} \]
(0.41) and the Euclidean $\ell_\infty$ norm is
\[ ||v||_\infty = \max_i |v_i|. \]
(0.42)

Theorem 0.77 (Equivalence of norms). Any two norms $||\cdot||_N$ and $||\cdot||_M$ on a finite dimensional vector space $V = \mathbb{C}^n$ satisfy
\[ \exists c_1, c_2 \in \mathbb{R}^+, \text{ s.t. } \forall x \in V, \quad c_1 ||x||_M \leq ||x||_N \leq c_2 ||x||_M. \]
(0.43)

Theorem 0.78. A function $||\cdot|| : V \to F$ is a norm iff it satisfies

- (NRM-1) real positivity: $\forall v \in V, ||v|| \geq 0$;
- (NRM-2) point separation: $||v|| = 0 \Rightarrow v = 0$.

- (NRM-3) absolute homogeneity: $\forall a \in F, \forall v \in V, ||av|| = |a||v||$;
- (NRM-4) triangle inequality: $\forall u, v \in V, ||u + v|| \leq ||u|| + ||v||$;

Remark 0.59. Theorem 0.78 is another common way to define a norm. The norm defined in (0.40) indeed satisfies conditions (NRM-1, 2, 3, 4). For (NRM-3),
\[ ||av||^2 = \langle av, av \rangle = a \langle v, av \rangle = a\bar{a} \langle v, v \rangle = |a|^2 ||v||^2. \]

To prove (NRM-4), we have
\[ ||u + v||^2 = \langle u + v, u + v \rangle = \langle u, u \rangle + \langle v, v \rangle + \langle u, v \rangle + \langle v, u \rangle \leq ||u||^2 + ||v||^2 + 2||u||||v|| = (||u|| + ||v||)^2, \]
where the third step uses Cauchy-Schwarz inequality Theorem 0.80 below.

Theorem 0.79 (Pythagorean). If $u, v$ are orthogonal, then
\[ ||u + v||^2 = ||u||^2 + ||v||^2. \]

Theorem 0.80 (Cauchy-Schwarz inequality).
\[ |\langle u, v \rangle| \leq ||u||||v||, \]
(0.44) where the inequality holds iff one of $u, v$ is a scalar multiple of the other.

Proof. For any complex number $\lambda$, (IP-1) implies
\[ \langle x_1 + \lambda x_2, x_1 + \lambda x_2 \rangle \geq 0 \Rightarrow \langle x_1, x_1 \rangle + \lambda \langle x_1, x_2 \rangle + \lambda \langle x_2, x_1 \rangle + \lambda^2 \langle x_2, x_2 \rangle \geq 0. \]
Set $\lambda = -\langle x_1, x_2 \rangle / \langle x_2, x_2 \rangle$, substitute into the above and we obtain (0.44).

Example 0.60. If $x_i, y_i \in \mathbb{R}$, then
\[ \sum_{i=1}^{n} x_iy_i \leq \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2. \]

Example 0.61. If $f, g : [-1, 1] \to \mathbb{R}$ are continuous, then
\[ \int_{-1}^{1} f(x)g(x)dx \leq \left( \int_{-1}^{1} f^2(x)dx \right)^{\frac{1}{2}} \left( \int_{-1}^{1} g^2(x)dx \right)^{\frac{1}{2}}. \]

Definition 0.81. In $\mathbb{C}^n$, a sequence of vectors $\{x_1, x_2, \ldots\}$ is said to converge to a vector $x$ iff $\lim_{m \to \infty} ||x_m - x|| = 0$.

0.8 Determinants

Definition 0.82. A permutation of the sequence $(1, 2, n)$ is a function that reorders this sequence. The set of all such permutations, denoted by $S_n$, is known as the symmetric group on $n$ elements. For each permutation $\sigma$, its signature, denoted by $\text{sgn}(\sigma)$, is $+1$ whenever the reordering given by $\sigma$ can be achieved by successively interchanging two entries an even number of times, and $-1$ whenever it can be achieved by an odd number of such interchanges.

Example 0.62. Suppose a sequence $(1, 2, 3)$ is reordered to $\sigma = (2, 3, 1)$. Then $\text{sgn}(\sigma) = -1$, $\sigma_1 = 2$, $\sigma_2 = 3$, and $\sigma_3 = 1$.

Definition 0.83 (Leibniz formula of determinants). The determinant of a square matrix $A \in \mathbb{R}^{n \times n}$ is
\[ \det A = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^{n} a_{i, \sigma_i}, \]
(0.45) where the sum is over all permutations $\sigma$ of the sequence $(1, 2, \ldots, n)$ and $a_{i, \sigma_i}$ is the element of $A$ at the $i$th row and the $\sigma_i$th column.

Definition 0.84. The $i,j$ cofactor of $A \in \mathbb{R}^{n \times n}$ is
\[ C_{ij} = (-1)^{i+j} M_{ij}, \]
(0.46) where $M_{ij}$ is the $i,j$ minor matrix of $A$, i.e. the determinant of the $(n-1) \times (n-1)$ matrix that results from deleting the $i$-th row and the $j$-th column of $A$.

Theorem 0.85 (Laplace formula of determinants). Given fixed indices $i, j \in \{1, 2, \ldots, n\}$, the determinant of an $n$-by-$n$ matrix $A = [a_{ij}]$ is given by
\[ \det A = \sum_{j'=1}^{n} a_{ij'} C_{ij'} = \sum_{i'=1}^{n} a_{i'j} C_{i'j}. \]
(0.47)